DOCUMENT RESUME

ED 356 211 SP 034 442

AUTHOR Vollmer, Marian L.; Creek, Roy J.
TITLE Teacher Assessment: A Continuing Controversy.
PUB DATE Feb 93
NOTE 13p.; Paper presented at the Annual Meeting of the American Association of Colleges for Teacher Education (San Diego, CA February 24-27, 1993).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Comparative Analysis; *Competence; Elementary Education; *Evaluation Criteria; Graduate Students; Higher Education; Measurement Objectives; *Measurement Techniques; Preservice Teacher Education; Student Teacher Evaluation; *Teacher Evaluation; Teacher Interns
IDENTIFIERS *NTE Specialty Area Tests; Pedagogical Content Knowledge; *Teacher Performance Assessment Instrument

ABSTRACT

Recent literature emphasizing the relationship of content knowledge to teaching performance has provided an inducement for some evaluators to use formal tests to assess teachers' understanding of subject matter. This study, conducted at a laboratory school connected with the University of Pittsburgh (Pennsylvania), examines the relationship between subjective and objective assessment instruments, specifically, the relationship between the Teacher Performance Assessment Instrument (TPAI), a measure focusing on teacher competence, and the National Teacher Examination's Education in the Elementary School Specialty Area Test (NTE/EES), covering knowledge in eight subject areas. Thirteen interns in the Master of Arts in Teaching program were evaluated by their assigned master teachers at the beginning of the Fall semester and toward the end of the Spring semester. The TPAI was used to assess their performance competencies and the NTE/EES was used to assess their subject knowledge. Data analysis revealed substantial gains on the TPAI as compared to moderate gains on the NTE/EES from one semester to the other. These data suggest that it would be possible to employ a teacher, based upon high NTE scores, who might receive poor evaluations based upon other performance instruments. One inference that can be drawn is that an intense practicum of sufficient duration produces statistically significant gains in both classroom performance and subject matter knowledge. (Contains 13 references.) (LL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED356211

TEACHER ASSESSMENT:

A Continuing Controversy

by

Marian L. Vollmer

and

Roy J. Creek

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

M Vollmer

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

SP034442

TEACHER ASSESSMENT: A CONTINUING CONTROVERSY

The assessment of teaching competence is a continuing educational concern (Davey, 1991). Teacher evaluation schemes ranging from the use of process rating scales to product assessments have had little impact upon the negative momentum of public opinion about the competence of classroom teachers, and evaluation practices continue to be held in low regard (Rebell, 1991). Today's researchers conclude that traditional assessment procedures have been naive, at times purposeless, and overly simplistic (Cruickshank & Haefele, 1990). Across the 16,000 school districts in the United States, the trend toward assessing teaching performance has increased considerably. As evaluation provisions are stipulated in negotiated contracts and school board policies, ways of measuring that competence will have to become less controversial than they have in the past.

Recent literature emphasizing the relationship of content knowledge to teaching performance has provided an inducement for some evaluators to use formal tests to assess teachers' understanding of subject matter. Proponents of this approach suggest that subject knowledge affects how teachers represent content to students (Grossman, 1991). For example, teachers with strong understandings of mathematics were more likely to encourage students to find different ways to solve problems and emphasize mathematical thinking. Teachers with weak understandings of the subject matter were more apt to view mathematics as a collection

of arbitrary algorithms and more likely to encourage students to merely find the right answer (Grossman, 1991). Although the standardized testing of subject matter has revealed a moderately strong relationship to success in teacher education programs, many researchers do not perceive the relationship to be valid for identifying the competencies needed for the day-to day performance of a classroom teacher (Haertel, 1987; Aye & Quells, 1979; Wittenberg & Weimar, 1973). Davey (1991) emphasizes this point, concluding that "teacher competence involves a complex set of knowledges, abilities, and personal attributes in dynamic interplay which cannot be captured by standardized paper and pencil tests. While certain levels of cognitive ability and basic knowledge, as assessed by standardized tests, are required for competence, there is more to the equation than these tests can realistically be expected to assess."

Initiatives toward identifying more sophisticated ways of evaluating the actual tasks which teachers must perform have originated with the National Board for Professional Teaching Standards, The Stanford Teacher Assessment Program, and the Connecticut Department of Education. The emphases of these projects are to identify the characteristics of effective teaching and develop assessment techniques that focus on performance based evaluation, interviews, and demonstration lessons (Rebell, 1991b). In these modes of assessment, the teacher is evaluated on involvement in the act of teaching and the responsiveness of others (e.g., students) instead of isolated behaviors (Weak and Evermore,

1991). Nevertheless, the National Teachers Examination continues to be widely used for purposes of certification and/or initial employment, which is why this project was undertaken.

The purpose of this investigation was to examine the relationship between the Teacher Performance Assessment Instrument (TPAI), a performance measure, and the National Teachers Examination (NTE), a content knowledge measure. The research questions, procedures, and an analysis of the data follow.

Research Questions

The goal of this study was to examine the relationship between a performance teacher assessment instrument, the TPAI, and a content knowledge instrument, the NTE. The population included 13 graduate interns enrolled in the Master of Arts in Teaching (MAT) Program in elementary education at the University of Pittsburgh. The MAT internship is a graduate program designed for college graduates with degrees in liberal arts, education, or other professional fields. Upon completion of the program, the students earn a Master of Arts in Teaching degree and are eligible for an Instructional I teaching certificate in Pennsylvania.

Procedures

Each MAT intern was evaluated by the master teacher to whom he/she was assigned. The Instrument used to assess performance competencies of the interns was the Teacher Performance Assessment Instruments (TPAI). The TPAI, which took four years to develop and

validate, was developed by the Georgia Teacher Assessment Project team. The TPAI focuses on teacher competence in the areas of: planning instruction and choosing materials to achieve instructional objectives; cognitive interaction with learners; skill in organizing and presenting instructional activities; and interpersonal skills associated with classroom climate and performance during instruction (Tanner & Ebers, 1985). Following the Conditions for Use of the TPAI, the pre-test (Fall Term assessment) was begun after the 20th school day (actual day of instruction), and was completed on or before the 70th school day. The post-test (Spring Term assessment) was begun after the 110th school day, with completion on or before the 160th school day (Capie, 1979, p. 20).

Using the same administration time line each MAT intern was evaluated on their teaching knowledge using the National Teacher Examination's Education in the Elementary School Specialty Area Test (NTE/EES). Eight subject areas are covered in the EES test: Language Arts, Reading, Mathematics, Science, Social Studies, Music, Art, and Physical Education and Health.

"According to the test specifications, every question should relate either to the child as the focus of teaching or to the process of teaching in the elementary-middle school...Every question is two dimensional. It is related to one of the eight subject matter areas and, at the same time, concerns one of the areas of knowledge about the child as the focus of teaching or the teaching process." (Educational Testing Service, ETS, 1989).

Analysis of the Data

A linear regression analysis was done to determine if pre-test scores on the NTE/EES were statistically related to the pre-test scores of the TPAI. An r = .127 revealed no correlation. Similarly, there was no correlation between NTE/EES and TPAI scores on the post-test, r = .063. In fact, an examination of Figures 1 and 2 reveals inverted extremes on both the pre-test and the post-test. One individual with a very high NTE score had a low TPAI score, while another subject who scored low on the NTE received a very high TPAI rating.

Figure 1. Pre-test relationship between TPAI and NTE scores.

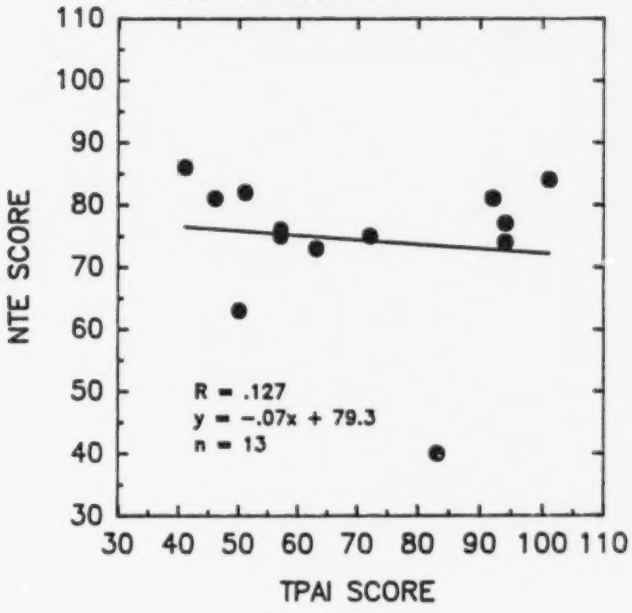
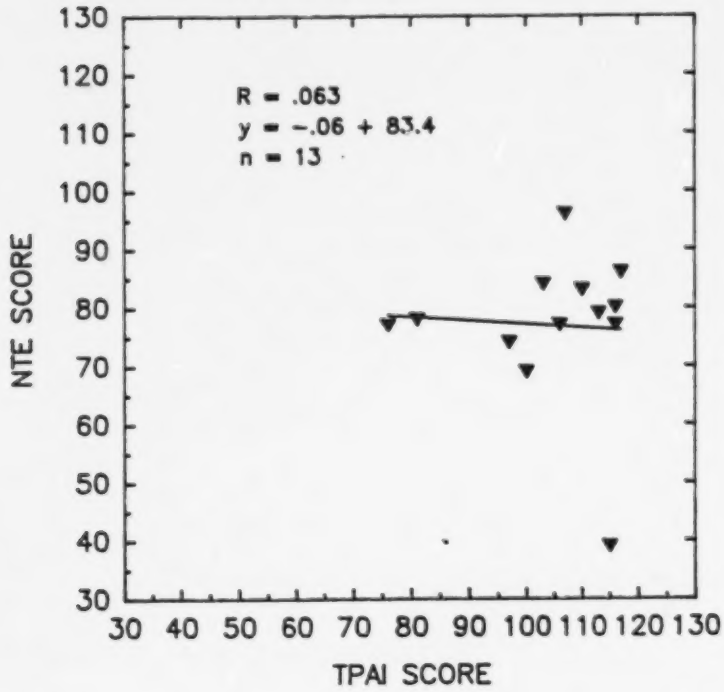


Figure 2. Post-test relationship between TPAI and NTE scores.



Nevertheless, it is interesting to see that the post-test scores were significantly higher that the pre-test scores on both measures. The mean pre- and post-TPAI scores were significantly different indicating a substantial improvement in overall competencies during the MAT experience (Table I). There was a smaller but significant improvement in NTE scores as well (Table II).

TABLE I. PRE- AND POST-TEST TPAI SCORES (Mean ± S.E.)

PRE-TEST	69.31 ± 5.8
POST-TEST	104.4 ± 3.7*
CHANGE	35.1 ± 5.1
n = 13	

* Significant difference from PRE-TEST, paired t-test, $t = 6.93$, $p < 0.001$.

TABLE II. PRE- AND POST-TEST NTE SCORES (Mean ± S.E.)

PRE-TEST	74.4 ± 3.3
POST-TEST	77.1 ± 3.6*
CHANGE	2.7 ± 1.1
n = 13	

* Significant difference from PRE-TEST, paired t-test, $t = 2.49$, $p < .05$.

Discussion

These results only begin to address the controversy surrounding teacher competency measures. At this time, 24 states require a passing grade on the NTE for entry into the teaching profession (Eissenberg and Rudner, 1988). Once in the profession, many teachers are evaluated using assessment instruments similar to the TPAI that rely upon observation, interviews, or other performance related variables. This data suggests that it is very possible to employ a teacher, based upon high NTE scores, who might receive poor evaluations based upon other performance type instruments. Because this peculiarity occurred on both the pre test and post test results, we are surmising that it is something more than a one-time aberration. If our further research reveals that it is a true discrepancy in the kinds of instrumentation used to evaluate teachers, the presumed relationship between subject matter knowledge and teaching performance will be called into question.

One inference that can be drawn from the data is that an intense practicum of sufficient duration produces statistically significant gains in both classroom performance and subject matter knowledge. Moreover, when the dramatic gains on the TPAI are compared to the moderate gains on the NTE, a whole new series of questions is generated. For example, does the MAT Program stress classroom performance more than content knowledge? Do students enter the program with substantial content knowledge and minimal

classroom performance skills? Is one of the instruments more objective or more accurate than the other? And most importantly, if the TPAI and the NTE are not assessing the same teaching competencies, which type of assessment tool should we use, and why?

References

Ayers, J., & Quall, G. (1979). Concurrent and predictive validity of the National Teachers Examination. Journal of Educational Research, 73, 86-91.

Capie, W., Johnson, C., Anderson, S., Ellett,C., & Okey, J. (1979). Teacher performance assessment instruments. Athens, Georgia: University of Georgia, 1979. (ERIC Document Reproduction Service No ED 183 518)

Cruickshank, D., & Haefele, D. (1990). Research-based indicators: Is the glass half-full or half-empty? Journal of Personnel Evaluation in Education, 3, 33-40.

Davey, B. (1991). Evaluating teacher competence through the use of performance assessment tasks: An overview. Journal Of Personnel Evaluation in Education, 5, 121.

Educational Testing Service. NTE A Guide to the Education in the Elementary Education School Specialty Area Test. Educational testing service, Princeton, N.J., 9.

Eissenberg T., & Rudner, L. (1988). State testing of teachers: A summary. Journal of Teacher Education, July/August, pp. 21-22.

Grossman, P. (1991). Mapping the terrain: Knowledge growth in teaching. In Effective teaching: Current research. edited by Waxman & Walberg. McCutchan Publishing Corp. 204-207.

Quirk, T., Witten, B., & Weinberg, S. (1973). Review of studies of the concurrent and predictive validity of the National Teacher Examination. Review of Educational Research, 43, 89-113.

Rebell, M.A. (1991). Teacher performance assessment: The changing state of the law. Journal of Personnel Evaluation in Education, 5, 227.

Shulman, L.S. (1987). Knowledge and teaching: Foundations of the new reform. Harvard Educational Review, 57(1), 1-22.

Tanner. C., & Ebers, S. (1985). Factors related to the beginning teacher's successful completion of a competency evaluation. Journal of Teacher Education, May-June, 41.

Weade, G., & Evertson, C. (1991). On what can be learned by observing teaching. Theory Into Practice. 30(1), 39.